

MAXIMUM MOMENTS OF SUM OF INDEPENDENT RANDOM MATRICES

MARCH T. BOEDIHARDJO

ABSTRACT. We show that the maximum moments of the sum of independent positive semi-definite random matrices with given norm upper bounds and norms of expectations is attained when all the random matrices are the multiplications of certain random variables and the identity matrix.

1. MAIN RESULT

Theorem 1.1. *Let $p, n, N \geq 1$ be natural numbers. Let $L_1, \dots, L_N > 0$ and $\alpha_1, \dots, \alpha_N \in [0, 1]$. Among all independent, positive semidefinite $n \times n$ random matrices X_1, \dots, X_N satisfying $\|X_k\| \leq L_k$ and $\|\mathbb{E}X_k\| = \alpha_k L_k$ for all $k = 1, \dots, N$, the quantity*

$$\mathbb{E} \circ \text{tr} \left(\sum_{k=1}^N X_k \right)^p$$

is maximized when $\mathbb{P}(X_k = L_k I) = \alpha_k$ and $\mathbb{P}(X_k = 0) = 1 - \alpha_k$.

In the case when $N = n$, $L_k = 1$ and $\alpha_k = \frac{1}{n}$, by [2, Theorem 2.5], it follows that for all independent, positive semidefinite $n \times n$ random matrices X_1, \dots, X_n satisfying $\|X_k\| \leq 1$ and $\|\mathbb{E}X_k\| = \frac{1}{n}$ for all $k = 1, \dots, n$,

$$\mathbb{E} \circ \text{tr} \left(\sum_{k=1}^n X_k \right)^p \leq \left(C \frac{p}{\log p} \right)^p n,$$

where C is a universal constant.

2. PROOF OF THE MAIN RESULT

Lemma 2.1 (Hölder inequality, [1], Corollary IV.2.6). *Let $p_1, \dots, p_r \geq 1$ with $\frac{1}{p_1} + \dots + \frac{1}{p_r} = 1$. Let A_1, \dots, A_r be $n \times n$ matrices. Then*

$$|A_1 \dots A_r|_1 \leq |A_1|_{p_1} \dots |A_r|_{p_r}$$

Lemma 2.2 (Araki-Lieb-Thirring inequality, [1], Exercise IX.2.11). *Let $\alpha \geq 1$. Let A, B be positive semidefinite $n \times n$ matrices. Then*

$$\text{tr}(ABA)^\alpha \leq \text{tr}(A^{2\alpha}B^\alpha).$$

As an immediate consequence, we have

Lemma 2.3. *Let $\alpha \geq 1$. Let A, B be positive semidefinite $n \times n$ matrices. Then*

$$|ABA|_\alpha \leq (\text{tr}(A^{2\alpha}B^\alpha))^{\frac{1}{\alpha}}.$$

Lemma 2.4. *Let $l_1, \dots, l_r, m_1, \dots, m_r \geq 1$. Let $l = l_1 + \dots + l_r$ and $m = m_1 + \dots + m_r$. Let X, Y be positive semidefinite $n \times n$ matrices. Then*

$$|\text{tr} X^{l_1} Y^{m_1} \dots X^{l_r} Y^{m_r}| \leq \|X\|^{l-1} \text{tr} X Y^m.$$

Proof.

$$\begin{aligned}
|\mathrm{tr} X^{l_1} Y^{m_1} X^{l_2} Y^{m_2} \dots X^{l_r} Y^{m_r}| &= |\mathrm{tr} X^{\frac{l_1}{2}} X^{\frac{l_1}{2}} Y^{m_1} X^{\frac{l_2}{2}} X^{\frac{l_2}{2}} Y^{m_2} \dots X^{\frac{l_r}{2}} X^{\frac{l_r}{2}} Y^{m_r}| \\
&= |\mathrm{tr} X^{\frac{l_1}{2}} Y^{m_1} X^{\frac{l_2}{2}} X^{\frac{l_2}{2}} Y^{m_2} \dots X^{\frac{l_r}{2}} X^{\frac{l_r}{2}} Y^{m_r} X^{\frac{l_1}{2}}| \\
&\leq |X^{\frac{l_1}{2}} Y^{m_1} X^{\frac{l_2}{2}} X^{\frac{l_2}{2}} Y^{m_2} \dots X^{\frac{l_r}{2}} X^{\frac{l_r}{2}} Y^{m_r} X^{\frac{l_1}{2}}|_1 \\
&\leq |X^{\frac{l_1}{2}} Y^{m_1} X^{\frac{l_2}{2}}|_{\frac{m_1}{m_1}} |X^{\frac{l_2}{2}} Y^{m_2} X^{\frac{l_3}{2}}|_{\frac{m_2}{m_2}} \dots |X^{\frac{l_r}{2}} Y^{m_r} X^{\frac{l_1}{2}}|_{\frac{m_r}{m_r}}.
\end{aligned}$$

To get the first equality, we write $X^{l_1} = X^{\frac{l_1}{2}} X^{\frac{l_1}{2}}$, $X^{l_2} = X^{\frac{l_2}{2}} X^{\frac{l_2}{2}}$ etc. The second equality follows by moving $X^{\frac{l_1}{2}}$ to the end of the the product in the trace. The last inequality follows from Lemma 2.1.

To simplify the notation, let $l_{r+1} = l_1$. Then we obtain

$$|\mathrm{tr} X^{m_1} Y^{m_1} \dots X^{l_r} Y^{m_r}| \leq \prod_{i=1}^r |X^{\frac{l_i}{2}} Y^{m_i} X^{\frac{l_i}{2}}|_{\frac{m_i}{m_i}}.$$

For each $1 \leq i \leq r$, we have

$$|X^{\frac{l_i}{2}} Y^{m_i} X^{\frac{l_i+1}{2}}|_{\frac{m_i}{m_i}} \leq \|X\|^{\frac{l_i}{2} - \frac{m_i}{2m} + \frac{l_i+1}{2} - \frac{m_i}{2m}} |X^{\frac{m_i}{2m}} Y^{m_i} X^{\frac{m_i}{2m}}|_{\frac{m_i}{m_i}} \leq \|X\|^{\frac{l_i+l_i+1}{2} - \frac{m_i}{m}} (\mathrm{tr} X Y^m)^{\frac{m_i}{m}},$$

where the first inequality follows from the ideal property of the Schatten norm and the second inequality follows from Lemma 2.3. Therefore,

$$\mathrm{tr} X^{l_1} Y^{m_1} \dots X^{l_r} Y^{m_r} \leq \prod_{i=1}^r \left(\|X\|^{\frac{l_i+l_i+1}{2} - \frac{m_i}{m}} (\mathrm{tr} X Y^m)^{\frac{m_i}{m}} \right).$$

The sum of the powers of $\|X\|$ is given by

$$\sum_{i=1}^r \left(\frac{l_i + l_{i+1}}{2} - \frac{m_i}{m} \right) = \frac{1}{2} \left(\sum_{i=1}^r l_i + \sum_{i=1}^r l_{i+1} \right) - 1 = \frac{1}{2} (l + l) - 1 = l - 1.$$

We conclude that

$$\mathrm{tr} X^{l_1} Y^{m_1} \dots X^{l_r} Y^{m_r} \leq \|X\|^{l-1} \mathrm{tr} X Y^m.$$

□

Lemma 2.5. *Let $l_1, \dots, l_r, m_1, \dots, m_r \geq 1$. Let $l = l_1 + \dots + l_r$ and $m = m_1 + \dots + m_r$. Let X, Y be independent, positive semidefinite $n \times n$ random matrices such that $\|X\| \leq L$. Then*

$$\mathbb{E} \circ \mathrm{tr} X^{l_1} Y^{m_1} \dots X^{l_r} Y^{m_r} \leq \mathbb{E} \circ \mathrm{tr} f^l Y^m,$$

where f is a random variable on $\{0, L\}$ independent from Y with $\mathbb{P}(f = L) = \frac{\|\mathbb{E} X\|}{L}$.

Proof. By Lemma 2.4,

$$\begin{aligned}
\mathbb{E} \circ \mathrm{tr} X^{l_1} Y^{m_1} \dots X^{l_r} Y^{m_r} &\leq \mathbb{E}(\|X\|^{l-1} \mathrm{tr} X Y^m) &\leq L^{l-1} (\mathbb{E} \circ \mathrm{tr} X Y^m) \\
&= L^{l-1} \mathrm{tr}(\mathbb{E} X)(\mathbb{E} Y^m) \\
&\leq L^{l-1} \|\mathbb{E} X\| \mathrm{tr}(\mathbb{E} Y^m) \\
&= \mathbb{E} f^l \mathbb{E} \circ \mathrm{tr} Y^m \\
&= \mathbb{E} \circ \mathrm{tr} f^l Y^m.
\end{aligned}$$

□

Lemma 2.6. *Let p be a natural number. Let X, Y be independent, positive semidefinite $n \times n$ random matrices such that $\|X\| \leq L$. Then*

$$(2.1) \quad \mathbb{E} \circ \text{tr}(X + Y)^p \leq \mathbb{E} \circ \text{tr}(fI + Y)^p,$$

where f is a random variable on $\{0, L\}$ independent from Y with $\mathbb{P}(f = L) = \frac{\|\mathbb{E}X\|}{L}$. Equality holds when X and Y commute, X/L is a projection, and $\mathbb{E}X$ is a scalar multiple of the identity.

Proof. Since $(X + Y)^p$ is a sum of products of the form $X^{l_1}Y^{m_1} \dots X^{l_r}Y^{m_r}$, (2.1) follows from Lemma 2.5. When X and Y commute,

$$\mathbb{E} \circ \text{tr}(X + Y)^p = \sum_{s=0}^p \binom{p}{s} \mathbb{E} \circ \text{tr} X^s Y^{p-s} = \sum_{s=0}^p \binom{p}{s} \text{tr}(\mathbb{E}X^s)(\mathbb{E}Y^{p-s}).$$

When X/L is a projection and $\mathbb{E}X$ is a scalar multiple of the identity, $\mathbb{E}X^s = (\mathbb{E}f^s)I$. Therefore,

$$\mathbb{E} \circ \text{tr}(X + Y)^p = \sum_{s=0}^p \binom{p}{s} \text{tr}(\mathbb{E}f^s)(\mathbb{E}Y^{p-s}) = \mathbb{E} \circ \text{tr}(fI + Y)^p.$$

□

Proof of Theorem 1.1. Let f_1, \dots, f_N be independent random variables on $\{0, L_1\}, \dots, \{0, L_N\}$, respectively, such that $\mathbb{P}(f_k = L_k) = \alpha_k$. We may assume that f_1, \dots, f_N are independent from X_1, \dots, X_N . Applying Lemma 2.6 recursively, we obtain

$$\begin{aligned} \mathbb{E} \circ \text{tr}(X_1 + \dots + X_N)^p &= \mathbb{E} \circ \text{tr}(X_1 + (X_2 + \dots + X_N))^p \\ &\leq \mathbb{E} \circ (f_1 I + (X_2 + \dots + X_N))^p \\ &= \mathbb{E} \circ \text{tr}(X_2 + (X_3 + \dots + X_N + f_1 I))^p \\ &\leq \mathbb{E} \circ \text{tr}(f_2 I + (X_3 + \dots + X_N + f_1 I))^p \\ &= \mathbb{E} \circ \text{tr}(X_3 + (X_4 + \dots + X_N + f_1 I + f_2 I))^p \\ &\leq \dots \\ &\leq \mathbb{E} \circ \text{tr}(f_1 I + \dots + f_N I)^p \\ &= \mathbb{E}(f_1 + \dots + f_N)^p, \end{aligned}$$

where the first inequality follows by taking $X = X_1$ and $Y = X_2 + \dots + X_N$ in Lemma 2.6 and the second inequality follows by taking $X = X_2$ and $Y = X_3 + \dots + X_N + f_1 I$.

Moreover, when X_1, \dots, X_N commute, X_k/L_k is a projection and $\mathbb{E}X_k$ is a scalar multiple of the identity, all inequalities become equalities. This happens, for instance, when $\mathbb{P}(X_k = L_k I) = \alpha_k$ and $\mathbb{P}(X_k = 0) = 1 - \alpha_k$. □

REFERENCES

- [1] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1997.
- [2] W. B. Johnson, G. Schechtman and J. Zinn, *Best constants in moment inequalities for linear combinations of independent and exchangeable random variables*, Ann. Probab. **13** (1985), 234-253.

DEPARTMENT OF MATHEMATICS, TEXAS A&M UNIVERSITY, COLLEGE STATION, TEXAS 77843
E-mail address: march@math.tamu.edu